

IT Standard 15.0.1: INTERNET DOCUMENT STANDARDS

Supersedes IT Standard 15.0.0: Internet Document Standards, issued 6/1/98

Software and Standards Management Branch
Systems Support Division
U.S. Census Bureau

Reissued November 24, 1999

Standard Number	Sequence: 15	Sub-standard: 0	Revision: 1
Category			
Key Words	HTML, relative link, PDF, URL, http, metadata		

Abstract

These guidelines and standards are established to ensure that all Census Bureau offerings on the Internet maintain a consistent corporate look and feel, and they preserve the Census Bureau's professional appearance and presence on the World Wide Web. Developing documents under these standards ensures that the system is user-friendly and accessible, with appropriate attribution for Census Bureau products and services.

Note: You can view this standard on the Intranet at http://cww.census.gov/tab/www/boc_guide.html. An annotation sheet is included with the HTML document. It contains information about anticipated changes that will be addressed during the next review of this standard. The changes are documented and posted so users are kept up-to-date until the next reissue. For further information regarding this listing on the Intranet contact Durwin Knutson at 301-457-1326 or Laura Yax at 457-2391.

Table of Contents

1	Scope	1
1.1	Reason for Standard	1
1.2	Who or What is Affected	1
2	Assumptions	1
3	HTML (Hypertext Markup Language) Document Standards	2
3.1	HTML Template	2
3.1.1	Levels of HTML Pages	2
3.1.2	Lower Level Template	2
3.2	Design Details for HTML Documents	3
3.2.1	World Wide Web (WWW) Page Documentation	3
3.2.2	Start of HTML Document	4
3.2.3	Head Section	4
3.2.4	Body Section	4
3.2.5	Background	5
3.2.6	Banner	5
3.2.7	Header	5
3.2.8	HTML Document Content	5
3.2.9	Compliance with the Privacy Act and Paperwork Reduction Act in Collections of Information on the Internet	9
3.2.10	Source and Contact Information	11
3.2.11	Footers	13
3.2.12	End of HTML Document	13
3.3	Other HTML Document Standards	13
3.3.1	Links	13
3.3.2	Image Tags	15
3.3.3	Transitional Pages	15
3.4	Standards for Other Documents	16
3.4.1	Other Mark Up Language Documents	16
3.4.2	Plain Unmarked Text (ASCII) Files	16
3.5	Adobe PDF (Portable Document Format) Files	17
3.6	Icons, Images, and Graphics Files	18
3.7	Audio and Video Files	19
3.7.1	Audio Files	19
3.7.2	Video Files	19
3.8	Definitions	19

4	Training	27
5	Reference Materials and Related Standards	27
	5.1 Reference Materials	27
	5.2 Related Standards	27
6	Conversion Requirements	28
7	Compliance Methods and Timetable	28
	7.1 Compliance Methods	28
	7.2 Timetable	29
8	Authority	29
	8.1 Policy	29
	8.2 Mandatory Standards	30
	Important Dates	31
	Contact for Additional Information	31

1 Scope

1.1 Reason for Standard

The purpose of this standard is to ensure that the Census Bureau's public Internet site maintains a consistent, easily recognized corporate look and feel and Bureau information served on the Internet site is easy to use and understand.

1.2 Who or What is Affected

This standard is mandatory for all persons authorized to write to the Census Bureau Internet servers. This includes Internet Coordinators and Data Disseminators as authorized by Form BC-1791A, Employee Authorization to Release or Change Information on the External Internet Site.

2 Assumptions

Assumptions that apply to this standard are as follows:

- users responsible for carrying out these standards have a working knowledge of Internet technology and practices;
- users affected by these standards understand the details of World Wide Web browsers and Hypertext Markup Language;
- HTML documents are the responsibility of the Data Disseminators and only these persons are authorized to make changes to the documents;
- the contents of all Internet documents are directly related to the mission and activities of the Census Bureau;
- all individuals working on the Internet have access to either workstations compliant with IT Standard 4.0.1: Workstation Configuration, or to desk-top personal computers compliant with IT Standard 7.0.5: Personal Computer (PC) Hardware Configuration.

3 HTML (Hypertext Markup Language) Document Standards

3.1 HTML Template

The template in this section contains the essence of the Census Bureau's corporate look and feel. The content of HTML documents *shall* directly relate to the mission and activities of the Census Bureau. Documents that promote individuals or contain staff pictures and biographical information are not allowed. Documents showing staff name or author and basic contact information are acceptable.

3.1.1 Levels of HTML Pages

Level 1 - This is the home page for the Census Bureau. There is only one Level 1 page; the Content Area of the Internet Management Team controls it.

Level 2 - There are a few Level 2 pages. They include the major topic areas listed on the Level 1 home page. Topical area pages at Level 2 are People, Housing, Business, Geography, and News. The divisions that produce their content maintain these pages. The template for Level 2 pages is available from the Internet Staff in Room 1364/3, phone 457-2157.

Lower Level - Any page that is not a Level 1 or Level 2 is a Lower Level page.

3.1.2 Lower Level Template

The author of a lower level HTML page *shall* use the following template:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 TRANSITIONAL//EN"
    "http://www.w3.org/TR/REC-html40/loose.dtd">
<HTML LANG="en">
<HEAD>
<TITLE> Your title here</TITLE>

<LINK REL="stylesheet" HREF="/main/.in/style.css" TYPE="text/css">
<META name="keywords" content="your, key, words, here, comma, separated">
<META name="description" content="your description">
</HEAD>

<BODY>
<A HREF="/"><IMG SRC="/main/www/img/cb_head.gif" BORDER="0" ALT="The U.S.
Census Bureau"></A>

<h2>header</h2>
```

body of document

Centering is encouraged, but not mandatory.

```
<CENTER>
<EM>Source: U.S. Census Bureau, division name,<BR>
branch name    </EM><BR>

Created:  mm, dd, yyyy <BR>

<!--#config timefmt="%B %d %Y" -->
Last revised:  <!--#echo var="LAST_MODIFIED"-->

</CENTER>
<P>
<!--#include virtual="/main/.in/footer.inc"-->
</BODY>
</HTML>
```

The template is found at: <http://cww.census.gov/tab/www/template.html>

3.2 Design Details for HTML Documents

3.2.1 World Wide Web (WWW) Page Documentation

HTML documents *should* comply with the most recent HTML standard. Currently this is HTML 4.0; the documentation is found on the W3C (World Wide Web Consortium [image of W3C]) site.

<http://www.w3.org/>



Pages *shall* be validated for proper HTML using the W3C validator service.

<http://validator.w3.org/>

If noncompliant standards are used for a HTML document, the author *shall* preview the document with an alternative browser to ensure readability and proper working links. A text-based browser such as Lynx *shall* be used for testing, and every effort *shall* be made to ensure that the document is readable in text-based browsers.

3.2.1.1 Page Accessibility

Authors of HTML documents *shall* create HTML documents according to the requirements of IT Standard 18.0.0: Design and Development of Web-Based Applications, to ensure their accessibility to individuals with disabilities.

3.2.2 Start of HTML Document

The following lines *shall* start every HTML document:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 TRANSITIONAL//EN"
    "http://www.w3.org/TR/REC-html40/loose.dtd">
<HTML LANG="en">
```

3.2.3 Head Section

The first element of the HTML document *shall* be the HEAD section.

```
<HEAD>
<TITLE> meaningful title is inserted here </TITLE>
<LINK REL="stylesheet" HREF="/main/.in/style.css" TYPE="text/css">
<META NAME="keywords" CONTENT="your, key, words, here, comma,
    separated">
<META name="description" content="your description">
</HEAD>
```

The TITLE appears inside the title bar of the window when the document is open in a web browser. It also appears on the list of results of a keyword search, so it *shall* be descriptive.

The TITLE tag and META tags “keywords” and “description” are used by the Census Bureau search engine. Searching is another way to aid users in finding information, including HTML and PDF documents, at the Census Bureau. Authors of HTML and PDF documents *shall* use these tags appropriately.

3.2.4 Body Section

The BODY section follows the HEAD section, and *shall* start with the following line:

```
<BODY>
```


3.2.5 Background

The HTML document background *shall* be white. The white background is designated, using its HEX code (#FFFFFF), in the cascading style sheet. Other colors or patterns *shall not* be used.

3.2.6 Banner

The first line of the BODY section *shall* include the official Census Bureau banner and *shall* be linked to the Census Home Page.

```
<A HREF="/" "><IMG ALT="U.S. Census Bureau" SRC="/main/www/img/cb_head.gif"
BORDER="0"></A>
```

This image *shall not* be copied or moved to another location. The author *shall* point to this image in /main/www/img.

There is only one Home Page at the Census Bureau Internet site and it is at:

<http://www.census.gov/>

No other page among the Census Bureau Internet offerings *shall* contain the words *home page* unless the author is referring to the Census Home Page or referring to home pages outside the Census Bureau.

3.2.7 Header

The next line of the BODY section *shall* be the document heading or a banner. If headings are used, the author *should* use the <H2> header style. Headers or banners are for organization and hierarchy, not for emphasis.

3.2.8 HTML Document Content

While the content of the BODY of the HTML document is determined by the author, standards of style have been developed to maintain a consistent look and feel, and a navigation scheme of the Census Bureau web site. In addition, it *shall* be the responsibility of each author to insure that all content on the Census Bureau web site meets the requirements of the Privacy Act and the Paperwork Reduction Act (see 3.2.9).

HTML documents *shall* be free of spelling and grammatical errors.

3.2.8.1 Cascading Style Sheets (CSS)

Authors *should* use CSS for layout and style where possible to aid in accessibility to the Census Bureau web site.

Distinguishing a document's structure and its presentation leads to more maintainable and reusable documents. Also, by extracting formatting directives from HTML documents, authors may design documents for a variety of users and target media in mind with minimal changes to their original HTML documents. The same HTML document, with different style sheets, may be tailored to color-blind users, those requiring large print, those with braille readers, speech synthesizers, hand-held devices, tty devices, etc. Furthermore, style sheets have another significant impact on accessibility. They eliminate the need to rely on "tricks" for achieving visual layout and formatting effects. These tricks have the unfortunate side-effect of making pages inaccessible.

For instance, HTML does not have an element or attribute to indent a paragraph, so many authors resort to using the BLOCKQUOTE element to indent text even when there is no quotation involved (many visual browsers indent the content of the element). This is misleading to nonvisual users: when an audio browser encounters a BLOCKQUOTE element, it should be able to assume that the enclosed text is a quotation. More often than not, that assumption proves incorrect since the element has been misused for a presentation effect.

The BLOCKQUOTE example shows the misuse, for presentation purposes, of an element intended to provide logical information. Many similar traps can seduce HTML authors: they use tables and invisible GIF images for the layout; they use H2 or H3 to change the font size of some text that is not a header; they use the EM element to italicize text when in fact, EM is meant to emphasize text (often presented with an italic font style, but rendered differently by a speech synthesizer); they use lists for alignment, etc.

CSS also allows users to override author styles. This is very important to users who cannot perceive a page with the author's chosen fonts and color. CSS allows users to view documents with their own preferred fonts, colors, etc. by specifying them in a user style sheet.

3.2.8.2 Fonts and Colors

Authors ***shall not*** alter colors of visited and non-visited links. This helps keep the Census Bureau web site easy to navigate and less confusing for users. The Census Home Page has been approved to change these colors. All other pages must use the defaults.

To enhance usability blue or purple colored text ***shall not*** be used because users are accustomed to those as default colors of hypertext links. Blue or purple colored unlinked text confuses navigation. Color blind users and users that access information using text-based browsers or screen readers are not aware of changes to text color. If you wish to emphasize text using color then please use one of the following HTML codes:

```
<STRONG CLASS="green">This text is in the color green.</STRONG>  
<STRONG CLASS="red">This text is red.</STRONG>
```

This will insure that our seeing users will notice the color change, and other browsers will also notice the change because of the tag that all browsers handle in some way that indicates a change to the tagged text. If you wish to emphasize in colors other than green or red, you will need to request the additional color through the Internet Staff.

Authors ***shall*** use the default font size for most of the page content.

3.2.8.3 Acronyms

Acronyms ***shall*** be explained. An author has three choices for explaining acronyms:

- provide a link to the Glossary of Acronyms at <http://www.census.gov/main/www/glossary.html>;
- provide a link from the acronym within the HTML document to its definition and full term or words;
- spell out the acronym followed by the abbreviation, for example, Minor Civil Division (MCD).

3.2.8.4 Length

HTML documents *should* be concise (maximum of 10 screens). Lengthy documents *should* use hypertext links to "name" anchors to aid navigation and enable access to detailed information. If all information fits on one screen, unnecessary links *should* be avoided.

3.2.8.5 Frames

The use of frames *shall* be avoided. Frames have a bug that allows content to be inserted from another server. Frames are a severe security risk. Also, frames are not widely accessible to all users. See 3.2.1.1 for information on page accessibility.

3.2.8.6 Tables

Tables are allowed but the author *shall* summarize the data and format the information in a way to make line by line reading sensible. The use of complex nested tables for page layouts *should* be avoided. The following example shows how to display information in a table format.

Example:

Ingredients	Cups	Ounces	Grams
Flour	1	4.4	125
Sugar	3 / 4	3.5	100
Butter	3 / 4	7	200

The table reads logically across the row. A speech synthesizer might render this table as: Ingredients, Cups, Ounces, Grams, Flour, 1, 4.4, 125, Sugar, 3/4, 3.5, 100, etc.

For the content to make sense the material should be kept as row-oriented as possible. Only the horizontal cells next to each other will be read in sequence; vertical cells probably will be separated by other cells.

3.2.8.7 Use of Staff Information

Documents *shall not* promote individuals or contain staff pictures and biographical information. Documents showing staff name, document author and basic contact information are permissible.

3.2.8.8 JavaScript

If JavaScript is used, there *shall* be a NOSCRIPT with text information to enable access to disabled users.

3.2.9 Compliance with the Privacy Act and Paperwork Reduction Act in Collections of Information on the Internet

Privacy Act Information:

Federal agencies must protect an individual's right to privacy when they collect personal information. This is required by the Privacy Act, 5 United States Code, 522a, and OMB Circular No. A-130, "Management of Federal Information Resources." OMB requires that every Federal web site include a privacy policy statement, even if the site does not collect any information that results in creating a Privacy Act record.

For systems of records where traditional paper collections of information are supplemented or replaced by electronic forms offered through a web site, the rules of the Privacy Act continue to apply. Thus, for situations where a Privacy Act notice is required in the paper-base world, the general principle is that the equivalent notice is required in the on-line world. Posting of the relevant Privacy Act notice on the web page or through a well-marked hyperlink would be appropriate.

Paperwork Reduction Act Information:

The Paperwork Reduction Act of 1995 (PRA), Public Law 104-13 (44 U.S.C. 3506 (c)(2)(A)), governs Federal collections of information from the public. The Office of Management and Budget (OMB) administers the PRA. OMB's implementing regulations are contained in 5 CFR Part 1320, "Controlling Paperwork Burdens on the Public."

The PRA defines a **collection of information** as "the obtaining, causing to be obtained, soliciting, or requiring the disclosure to third parties or the public, of facts or opinions by or for an agency, regardless of form or format, calling for either -- answers to identical

questions posed to, or identical reporting or recordkeeping requirements imposed on, ten or more persons, other than agencies, instrumentalities, or employees of the United States; or answers to questions posed to agencies, instrumentalities, or employees of the United States which are to be used for general statistical purposes . . .”

In other words, a collection of information is asking identical questions of more than nine people. Federal employees do not count unless the information collected is used for general statistical purposes. A collection of information may take many forms, for instance, a paper questionnaire, a personal or phone interview, or any of a variety of electronic questionnaire formats, including questions posed via the Internet.

Asking an Internet user to provide anything more detailed than his or her name or other basic identifying information such as an e-mail address requires OMB clearance as a collection of information. An undirected request for user comments, questions, complaints, or problems such as is typically collected through a “suggestion box” is considered exempt from the PRA.

If a collection of information is subject to the PRA, it must:

- have current valid OMB approval,
- display a current valid OMB control number and expiration date on the introductory screen,
- inform the respondent about
 - reason(s) for the collection,
 - use(s) of the information,
 - time estimates for response, along with a request for comments on the estimate,
 - whether responses are voluntary, mandatory (citing mandatory authority), or required to obtain or retain a benefit,
 - nature and extent of confidentiality provided to responses (citing authority, if any),
 - the fact that the respondent is not required to respond to, and an agency may not conduct any collection of information that is not approved by OMB.

If authors of HTML documents have questions about whether their planned Internet information collection needs OMB clearance, or for information on the clearance process, contact Tom Smith, Forms Clearance Officer at 301-457-3152.

A Comparison of the Required Disclosures to Survey and Census Respondents Under the Paperwork Reduction Act (PRA) and the Privacy Act (PA)

DISCLOSURE	PRA	PA
Display a currently valid OMB control number and expiration date on the survey instrument	Yes.	No.
State the reason(s) for the collection	Yes.	Yes.
State the use(s) that will be made of the information collected	Yes.	Yes. Must tell about routine uses made within the Agency and by other Agencies to whom personally identifiable information is disclosed.
Give a time estimate for response, along with a request for comments on that estimate	Yes.	No.
Tell whether responses are voluntary, mandatory, or required to obtain or retain a benefit	Yes. Must cite authority for mandatory collections.	Yes. Must cite authority for voluntary and mandatory collections. Must also specify the effects on the respondent of not providing all or any part of the information requested.
State the nature and extent of confidentiality provided to responses (citing authority, if any)	Yes.	No.
State the fact that the respondent is not required to respond to, and an agency may not conduct any collection of information that is not approved by OMB	Yes.	No.

3.2.10 Source and Contact Information

All Census Bureau HTML documents *shall* close with an agency identifier, division, creation date (if appropriate), and revision date:

```
<CENTER>
Source:  U.S. Census Bureau, Division, (Optional: Branch)
Created:  month day, year (e.g., Created:  March 25, 1999)
Last revised:  month day, year (e.g.,  Last revised: May 25, 1999)
</CENTER>
```

The agency identifier is: U.S. Census Bureau

This identifier must be on every page produced by the Census Bureau.

Division is the name of your division, such as “Housing and Household Economic Statistics Division,” or “Population Division.” Do not use abbreviations. Your division identifier must be on all pages that relate to data/information produced by your division.

A creation date is the date that the page is released to the public, not the date that the Public Information Office places the page on their embargo site. Creation date does not apply to pages that help route users to specific pages, such as the pages for “Disability,” “Poverty,” or “Education.” Creation date does apply to pages that contain specific reports and tables for a specific topic, such as a “Disability - Table 1 for 1999”.

Revision date applies to every page. Even if a minor cosmetic change is made, the date should change to reflect that a change occurred. This tells users that we keep our pages up-to-date. The last revision date may be generated by a call to the server to get the date stamp from a file. Use the following two lines of HTML code to generate this date stamp:

```
<!--#config timefmt="%B %d %Y" -->  
Last revised: <!--#echo var="LAST_MODIFIED"-->
```

Output is: Last revised: March 25, 1999

Optional lines of source information *should* be provided.

Author: the author of the HTML document
Contact: E-mail contact

The above items of information can be included using an include file. For example:

Create a file in your directory: /division/.in/person.inc

This include file will only include the following text:

```
Maintained By: John X. Smith (Something Division)<BR>
```

In your HTML file you can include this information by inserting the following text:

```
<!--#include virtual="/division/.in/person.inc"-->
```

This is an easy way to include text that will change, such as contact information.

3.2.11 Footers

All HTML documents *shall* include the Census Bureau footers.

```
<!--#include virtual="/main/.in/footer.inc"-->
```

Output:

```
<CENTER><hr width="60%"><FONT FACE="Helvetica" SIZE="1">
<a href="/population/www/people.html" target="_top">People</a>
| <a href="/hhes/www/housing.html" target="_top">Housing</a>
| <a href="/econ/www/index.html" target="_top">Business</a>
| <a href="/geo/www/index.html" target="_top">Geography</a>
| <a href="/pubinfo/www/news.html" target="_top">News</a>
| <a href="/main/www/subjects.html" target="_top">Subjects A to Z</a>
| <a href="/main/www/srchtool.html" target="_top">Search</a>
| <a href="/mp/www/censtore.html" target="_top">Catalog</a>
| <a href="/main/www/access.html" target="_top">Data Tools</a>
| <a href="/index.html" target="_top">Census Home</a>
</FONT> </CENTER>
```

Any additional subject footers *shall* be placed before the Census footers.

```
<!--#include virtual="/population/www/.in/popfoot.inc"-->
<!--#include virtual="/main/.in/footer.inc"-->
```

3.2.12 End of HTML Document

The last lines of the HTML code *shall* be

```
</BODY>
</HTML>
```

3.3 Other HTML Document Standards

3.3.1 Links

URLs (Uniform Resource Locator) within hypertext links *shall* be enclosed in "double quotations."

```
<A HREF="/hhes/www/poverty.html">Poverty</A>
```

URL names *shall* use appropriate case. UNIX is case sensitive with a preference for lowercase at the Census Bureau. To make links more accessible, use text that makes

sense when read out of graphic context. For instance instead of “to access click here” use “access The American FactFinder.” Refer to 3.2.1.1 for information on page accessibility.

Authors of HTML documents **shall** use relative links when linking documents on the same server. Links **shall** be relative to the document root (/) for documents outside the author’s directories. Links for documents inside the author’s directories **shall** be relative to that author’s directories.

For example:

The relative link to Estimates from the Census home page is:

```
<A HREF="/population/www/estimates/popest.html">Estimates</A>
```

The relative link to Estimates from the People page is:

```
<A HREF="./estimates/popest.html">Estimates</A>
```

Authors of HTML documents are responsible for ensuring that the links are accurate. Links **shall** be verified, particularly those connecting to outside resources.

Links to other sites **should** be limited to subject related links, Federal agencies, partnerships, State Data Centers, and nonprofit organizations. Links to commercial sites **should** be limited to

- resources that are essential to using Census Bureau materials, such as the Adobe Reader;
- secondary data suppliers;
- resources associated with clearinghouse lists, such as TIGER vendor and data repackager lists maintained by the Geography Division.

Links to sites outside the Census Bureau generally **should** point to home pages. Links pointing to pages below the home page level are subject to change and can make Census Bureau HTML documents appear unstable. All external links **should** contain the TITLE tag with the following text:

```
<A HREF="http://outside.census.site/" TITLE="This external site is not  
endorsed by the US Census Bureau">Someplace Else</A>
```

HTML documents **should not** contain dead links. A link to a temporary document is acceptable for no longer than 30 days. The temporary document **shall**

- have the corporate look and feel,
- identify the anticipated data product/service,
- provide the expected date of issue.

3.3.2 Image Tags

When using graphics or charts, an IMG ALT="description" **shall** be included. This applies to all graphics and visuals. A bullet icon is represented by ALT="o" or ALT="*." If the graphic itself contains much information, a summary or link to a longer description **shall** be provided in addition to the ALT="description." Wherever possible, text versions **should** be provided for all graphics screens. See 3.6 for specific instructions on graphics, images, and icons.

3.3.3 Transitional Pages

If the URL of a document changes, the author **should** provide a *transitional* link to the new location. The transitional link **should** be maintained for three months.

Before an existing page is removed, the author **should** use the "Link Finder" tool found on the ICG IntRAnet Page: <http://cww.census.gov/tab/www/icg.htm>. The author is responsible for notifying page authors of the change before removing the page.

A page can automatically load another page with a META REFRESH tag stated in the HEAD section of the HTML.

Example:

```
<META HTTP-EQUIV="Refresh" CONTENT="5;URL=new.html">
```

Usage: Content="N;URL=url"

N is the number of seconds to wait before loading specified page
url is the new URL to load

3.4 Standards for Other Documents

3.4.1 Other Mark Up Language Documents

Other mark up language files *shall* follow the intentions of the standards as written in 3.2 and 3.3.

3.4.2 Plain Unmarked Text (ASCII) Files

3.4.2.1 Text Files for Browser Display

Text documents for browser display *should not* exceed 80 characters in width and *shall not* exceed 120 characters in width. Each document *shall* include the phrase "Source: U.S. Census Bureau."

Additional metadata that includes documentation of file and record descriptions *shall* either be included in each document or have a reference to a separate document that contains this metadata. This additional metadata *should* at a minimum provide the following:

- Consistent with. . . [title of a report, a survey, etc.],
- Release date of the information,
- Date of last revision.

3.4.2.2 Text Files for Downloading

Data files for downloading *should* be available in ASCII text or delimited ASCII. Other widely used formats such as SAS transport, spreadsheet, and database formats *should* also be provided, as appropriate.

Additional metadata that includes documentation of file and record descriptions *shall* be included in a separate document. This additional metadata at a minimum *should* provide the following:

- Consistent with. . . [title of a report, a survey, etc.],
- Release date of the information,
- Date of last revision.

Authors *shall* give file size listings and descriptions. Large files more than 1.5 MB *should* be provided in compressed form for quicker downloading.

Compressed files *shall* list both the compressed and uncompressed file sizes. Files more than 100 MB are better suited to a different distribution method, such as CD-ROM or other application software.

The following compressing practices *shall* apply to data files for downloading:

- archive data sets then compress the data set into one file,
- offer data sets as individual files and as one compressed file,
- compress files using standard compression formats (e.g., pkzip, winzip, compress, and gzip),
- avoid self-extracting (DOS-only) compression files.

3.5 Adobe PDF (Portable Document Format) Files

Internet documents leading to PDF files *shall* include the following line on the page referring to the PDF document:

```
<!--#include virtual="/main/.in/pdf.inc"-->
```

Output:

```
The symbol <A HREF="/main/www/pdf.html">
<IMG SRC="/main/www/img/pdficon.gif" ALT="PDF" BORDER="0"></A> or
the letters PDF indicate documents in the
<A HREF="/main/www/pdf.html">Portable Document Format (PDF)</A>.
In order to view these files, you will need the
<A HREF="/main/www/pdf.html#reader">Adobe(R) Acrobat(R) Reader</A>
which is available for free from the Adobe web site.
```

There *shall* be an indication next to the filename that the document is in PDF, either by using the icon or the letters PDF. The file size *shall* also be stated. When PDF files are created, they *shall* be optimized and embedded graphics *shall* be minimized to reduce file size.

ASCII or HTML files *should* be available as an alternative to PDF files.

3.6 Icons, Images, and Graphics Files

3.6.1 A library of approved Census Bureau graphics and icons is listed at:

<http://cwww.census.gov/tab/www/library.html>

Authors **shall not** use navigation buttons that duplicate links in footers. Authors **shall not** change, copy, or resize graphics found in the graphics library.

If a graphic icon is used in the Internet document design, the author **shall** check the library to see if there is an existing graphic/image/icon. If there is an existing graphic/image/icon in: /main/www/img authors **shall** use the available image. Graphics/Images/Icons **shall not** be copied or moved to any other directory. If none is available, the user may create one. Graphics **shall** only contain information that is relevant and appropriate to the Census Bureau. The Graphics and Creative Design Section in ACSD will create graphics upon request.

3.6.2 All graphics showing data from the Census Bureau **shall** include attribution, e.g., Source: U.S. Census Bureau

3.6.3 Image files **shall** be in gif or jpeg (Joint Photographic Experts Group) format. Page graphics **should not** total more than 50 KB.

3.6.4 In-line graphics files **should not** be more than 10 KB in size. The author **should** closely crop gif images and reduce the amount of colors as low as possible to produce a good image. A site to help you reduce your graphic size without losing quality is found at:

<http://www.netmechanic.com/accelerate.htm>

3.6.5 The use of image maps **should** be avoided. When an image map must be used, the author **should** use client-side MAP and include text for image map zones to make the site accessible. In addition, links reached by image maps cannot be located by the site's search software. See 3.2.1.1 for information on page accessibility.

3.6.6 Color resolution and compatibility with background colors **should** be a consideration during the design phase. Graphics **shall** be transparent. The author **shall** create images using a web-safe color palette. Choosing colors from the web-safe palette

makes images less likely to look dithered or blotchy. More information about web-safe color palettes can be found at:

<http://www.lynda.com/hex.html>

3.7 Audio and Video Files

Internet documents that include audio or video files *shall* explicitly note the format and size (or running time for streaming audio) of the file. Such documents *shall* also include links to sources for downloading the software needed to play these files.

3.7.1 Audio Files

Audio files *shall* be provided in the most common, multi platform format(s) possible, such as Wave (.wav) and RealAudio (.ra). Captioning or transcripts *shall* be provided for all audio content.

3.7.2 Video Files

Video files *shall* be provided in the most common, multi platform format(s) possible, such as RealPlayer. Descriptions and transcripts *shall* be provided for all video content. Credit to the U.S. Census Bureau *shall* be included.

3.8 Definitions

3.8.1 absolute link: A hypertext reference to another document that specifies the host and domain. The opposite is a relative link. [Team Definition].

Example: <http://www.census.gov/prod/www/titles.html>

3.8.2 accessibility: Making sure that individuals with disabilities have access to and use of information and data that is comparable to the access to and use of the information and data by such individuals who are not with disabilities. [PL 105-220, 1998 HR 1385 PL 105-220, enacted on August 7, 1998, 112 Stat 936 codified as: Section 504 of the Rehabilitation Act, 29 U.S.C. § 794d, Section (A) DEVELOPMENT, PROCUREMENT, MAINTENANCE, OR USE OF ELECTRONIC AND INFORMATION TECHNOLOGY.]

3.8.3 ALT: In HTML, ALT is an attribute to the tag. The ALT attribute defines a text string to be put in place of the image for text-only browsers. Also see IMG. [Morris, Mary E. S. HTML for Fun and Profit. Sunsoft Press, 1995]

3.8.4 anchor: In HTML, an anchor can be a pointer to (1) a different document (see link) or (2) a specific place within a document. This kind of anchor is also called a *Named Anchor*. The anchor is named by using the NAME attribute with the <A> tag. For example: Subjects Starting With D. To reference this specific point, the link uses a # sign to show the named pointer within the document. For example: D [Team Definition]

3.8.5 ASCII: Abbreviation for American Standard Code for Information Interchange. ASCII is the Census Bureau's preferred binary encoding of alphabetical characters, numbers, and other keyboard symbols. It is the predominant character set encoding of present-day computers. [Team Definition]

3.8.6 audio file: A file with digital encoding that produces sound. The encoding may be linear, logarithmic, or mu-law. Sound is one component of multimedia. Filename extension: .au (Audio), .ra (Real Audio), .wav (WAVEform audio), and .aux (Windows). [Team Definition]

3.8.7 BODY: In HTML, a tag starting with <BODY> followed by the text (content) of the document and ending with </BODY>. [Team Definition]

3.8.8 browser: A program or application which communicates with a web server using the hypertext transfer protocol to send requests and receive responses. The application locates a document through a URL, retrieves it, and formats it for display to the user. Common browsers for the World Wide Web include:

- a. text browsers - which only displays text; no icons/images are displayed (e.g., Lynx), and
- b. graphic browsers - displays icons/images as well as text (e.g., Netscape, MSExplorer). [Team Definition]

3.8.9 compression: A procedure to make large files (more than 1.5 MB) smaller for efficient storage and for faster transfer across the Internet. Different computer systems use different compression software. [Team Definition]

3.8.10 corporate look and feel: As interpreted at the Bureau of the Census, it is the consistent appearance of **ALL** Internet documents. It is characterized by a system of

stylistic, informational, and navigational features that provide a subtle reminder to users that they are using information from the Bureau of the Census. [Team Definition]

3.8.11 Data Disseminators: Persons with write privileges to the Census Bureau Internet site. [Team Definition]

3.8.12 directory: A directory is a special type of file that contains a list of subordinate directory names and file names, as well as information on the size and date of the listed directories and files. [IT Standard 5.0.0]

3.8.13 directory pathname: The complete name by which the directory is known. The pathname gives the sequence of directories by which the directory is linked to the root directory. [Morris, Mary E. S. HTML for Fun and Profit. Sunsoft Press, 1995]

3.8.14 document: A document as applied to the Census Bureau's Internet site is any kind of file being served on the Internet site. Each document/file has a unique URL. [Team Definition]

3.8.15 extension: In reference to file names, an extension is a set of characters added to a file name that serves to define the type of file. The extension is usually the characters that follow the period in a file name (e.g., .txt, .pdf, .html). [Team Definition]

3.8.16 file: A named collection of data (information) stored on a disk or tape media. [IT Standard 5.0.0]

3.8.17 filename: The name of a file as it is stored in a directory on a disk. Also see pathname. [Morris, Mary E. S. HTML for Fun and Profit. Sunsoft Press, 1995]

3.8.18 FTP (File Transfer Protocol): This is a service; programs and a set of specifications that provide the ability to copy a file from any Internet host to any other Internet host. At the Census Bureau, FTP is the action of transferring a file to a computer. FTP is also used to describe directories on the Census Bureau's Internet sites that contain large files suitable for transferring (e.g., compressed files). [Team Definition]

3.8.19 gif (Graphics Interchange Format): A standard for digitized images. A gif is a common way of encoding an image. Filename extension: .gif [Team Definition]

3.8.20 HEAD: In HTML, a tag starting with <HEAD> followed by the title and other appropriate items for the head section of a HTML document. The section ends with </HEAD>. [Team Definition]

3.8.21 Home Page: The top-level document relating to an individual or institution. This usually has a URL consisting of a host name. The host is the name of the computer serving the information. There is only one Home Page at the Census Bureau; it is found at <http://www.census.gov/>. All other documents on a server *should* be accessible by following links from the Home Page. [Team Definition]

3.8.22 horizontal rule: In HTML, a horizontal line is displayed when either the <HR> tag is processed or a gif image of a line is displayed. Horizontal rule is used to divide sections of a document. [Team Definition]

3.8.23 HREF (hypertext reference): In HTML, the address of the destination of a hypertext link. In WWW, hypertext references are given in URL format. [Morris, Mary E. S. HTML for Fun and Profit. Sunsoft Press, 1995]

3.8.24 HTML (Hypertext Markup Language): The document-formatting language used by WWW browsers. [Morris, Mary E. S. HTML for Fun and Profit. Sunsoft Press, 1995]

3.8.25 http (Hypertext Transfer Protocol): The Internet protocol used to transfer documents and other MIME data between systems. The Hypertext Transfer Protocol (HTTP) is an application-level protocol for distributed, collaborative, hypermedia information systems. HTTP has been in use by the World-Wide Web global information initiative since 1990. HTTP allows basic hypermedia access to resources available from diverse applications, such as FTP, and others. [Morris, Mary E. S. HTML for Fun and Profit. Sunsoft Press, 1995, HTTP 1.1 Internet Draft]

3.8.26 hypertext: A document that contains links to other documents, indicated by text in a special format. The linked document can be actively followed or referenced by activating the emphasized text. [Morris, Mary E. S. HTML for Fun and Profit. Sunsoft Press, 1995]

3.8.27 icon: An on-screen symbol that simplifies access to a program, command, or data file. An icon is a small picture intended to represent something (a file, directory or action) in a graphical user interface. When an icon is clicked, some action is performed such as opening a directory or aborting a file transfer. Icons are usually stored as bitmaps. [Team Definition]

3.8.28 image: A picture or graphic representation of a two-dimensional scene. It is composed of pixels arranged in a rectangular array. Images are usually taken from the real world via a digital camera, frame grabber or scanner, or created by certain types of software. They are usually stored as gif or jpg files. [Team Definition]

3.8.29 image map: An image in a HTML document with "hot spots" which, when activated (in a suitable browser), acts as anchors or links to other information. The Bureau's data map at <http://www.census.gov/datamap/www/index.html> is an example. [Team Definition]

3.8.30 image map zones: These are the defined, clickable areas of an image map, also called "hot spots". [Team Definition]

3.8.31 IMG: The IMG element embeds an image in the current HTML document at the location of the element's definition. The IMG element has no content; it is usually replaced in line by the image designated by the src attribute. [HTML 4.0 Standard]

3.8.32 internet: A collection of networks interconnected by a set of routers that enable them to function as a single, large virtual network. [Morris, Mary E. S. HTML for Fun and Profit. Sunsoft Press, 1995]

3.8.33 Internet: The largest internet in the world consisting of large national backbone nets (such as MILNET, NSFNET, and CERN) and a myriad of regional and local campus networks all over the world. The Internet uses the Internet protocol suite. To be on the Internet, the user must have IP connectivity and communicate via TCP/IP protocols. Networks with only E-mail connectivity are not actually classified as being on the Internet. [Morris, Mary E. S. HTML for Fun and Profit. Sunsoft Press, 1995]

3.8.34 Intranet: An internet internal to an agency. Parts of these nets are sometimes called the Corporate Wide Web (CWW). The contents of an Intranet are usually not accessible to persons outside the agency. [Team Definition]

3.8.35 jpeg: Acronym for Joint Photographic Experts Group.

3.8.36 Joint Photographic Experts Group (jpeg): jpeg is an image compression algorithm used to store photographs and images on computer networks. File name extension: .jpg or .jpeg. [Team Definition]

3.8.37 last revision date: This date refers to the date and time that any change was made to a static HTML document on the Internet. [Team Definition]

3.8.38 link: A hypertext reference to another document or another point in the same document. [Morris, Mary E. S. HTML for Fun and Profit. Sunsoft Press, 1995]

3.8.39 metadata: Data about data. Data definitions describing aspects of actual data items (in a document) such as a name, the source, format. [Team Definition]

3.8.40 Multipurpose Internet Mail Extensions (MIME): A standard for multiple part, multimedia electronic mail messages and World Wide Web hypertext documents on the Internet. MIME provides the ability to transfer nontextual data, such as graphics, audio and fax. [Team Definition]

3.8.41 page: A page, in WWW parlance, is an individual computer file or document which can be addressed by a hypertext link. This usually means a HTML document. Although most pages are static files, some are not. HTML forms, Common Gateway Interface (CGI) scripts, and other mechanisms may be used to create pages which are front-ends to databases or other types of information services. Such services may dynamically create new pages containing query responses or customized forms in response to user input. [Team Definition]

3.8.42 pathname: The location of a file or directory in a file system. [Morris, Mary E. S. HTML for Fun and Profit. Sunsoft Press, 1995]

3.8.43 PDF (Portable Document Format): The native file format for Adobe Systems' Acrobat. PDF is the file format for representing documents in a manner that is independent of the original application software, hardware, and operating system used to create those documents. A PDF file can include any combination of text, graphics, and images in a device-independent and resolution independent format. PDF documents can be one page or thousands of pages, very simple or extremely complex with a rich use of fonts, graphics, color and images. ["The Portable Document Format Reference Manual," Adobe systems Inc. Addison-Wesley Publ. Co., ISBN: 0-201-62628-4]

3.8.44 relative link: A hypertext reference to another document that uses a path relative to the source directory. The opposite is an absolute link. [Team Definition]

Example: /prod/www/titles.html

3.8.45 search engine: Any number of software packages that allow users to specify a keyword or a character string. The software searches the contents of the Internet (usually the WWW) to find documents with occurrences of the specified characters. Each software package uses a different approach to locating Internet offerings. Some examples of search engines are Alta Vista, Lycos, and Infoseek. [Team Definition]

3.8.46 server (web server): An application which provides access to one or more collections of documents using the hyper text transfer protocol. A web server accepts connections in order to service requests by sending back responses. [Team Definition]

3.8.47 shall: The word *shall* indicates a mandatory statement. [IT Standards 1.0.2]

3.8.48 should: The word *should* indicates a guideline which is recommended, but not mandatory. [IT Standard 1.0.2]

3.8.49 source: This is the agency/department that produced the data/report being posted to the internet.

3.8.50 TAG: In HTML, a sequence of characters starting with a < and ending with a >, used to mark up text with display instructions. [Morris, Mary E. S. HTML for Fun and Profit. Sunsoft Press, 1995]

3.8.51 TITLE: In HTML, a tag starting with <TITLE> followed by a title/description of the contents of the page and ending with </TITLE>. The title *should* be as short as possible but fully informative and specific. The contents of the title are displayed in the top area of most web browsers. Although the title is often overlooked because it does not appear in the body of the document, it is important because it is frequently used to identify the document on hot lists, search result sets, and site indexes. [Team Definition]

3.8.52 transitional link: A temporary link that *shall* be used when a URL has been moved from one address to another. The Census Bureau has determined that transitional links *shall* be maintained for three months. [Team Definition]

3.8.53 URL (Uniform Resource Locator): (Previously known as Universal Resource Location) A draft standard for specifying an object on the Internet, such as a file or news group. URLs are used extensively on the World Wide Web. They are used in HTML documents to specify the target of a hyperlink. Examples of URLs include:

ftp://wuarchive.wustl.edu/mirrors/msdos/graphics/gifkit.zip
http://www.w3.org/default.html

news:alt.hypertext
telnet://dra.com
mailto:dbh@doc.ic.ac.uk
<http://wombat.doc.ic.ac.uk/?Uniform+Resource+Locator>
<http://www.w3.org/default.html#Introduction>

How to read a URL: The part before the colon specifies the access scheme or protocol used by the specific site (e.g., http, telnet, ftp.) The part after the double slash is the location or host name of the computer being accessed (e.g., www.census.gov, www.whitehouse.gov, etc.) The part after the host name references the directory structure where the resource is stored. The URL may just reference a directory (e.g., <http://www.census.gov/population/www/>) or it may reference a specific file in a directory (e.g., <http://www.census.gov/datamap/www/26.html>).

Census Bureau rules state that relative URLs *shall* be used whenever possible, for both anchors and images, in order to make documents and collections portable. URL names *shall* use lowercase. Arbitrary use of uppercase (e.g., picture.GIF, myFILE.html) is prohibited. [Team Definition]

3.8.54 webmaster: Each WWW server has a webmaster--an individual with primary responsibility for the server. By convention, each WWW server maintains a Webmaster E-mail address (e.g., webmaster@census.gov). At the Census Bureau, webmaster goes to the Internet staff. [Team Definition]

3.8.55 World Wide Web (WWW): A network document publishing system that uses Hypertext Markup Language (HTML) as the authoring language and Hypertext Transport Protocol (HTTP) as the transport protocol. There are a variety of client and server products that support HTML and HTTP. See also Hypertext Markup Language and Hypertext Transfer Protocol. [Morris, Mary E. S. HTML for Fun and Profit. Sunsoft Press, 1995]

3.8.56 W3C (World Wide Web Consortium): The main standards body for the World Wide Web. W3C works with the global community to produce specifications and reference software. [W3C Home Page at <http://w3.org/pub/www/The Project.html>]

4 Training

Training is required for all Internet Coordinators and Data Disseminators. The Census Bureau's Internet Staff and Internet Coordinators Group provides training which includes a

segment on the requirements contained in this standard. Classes are scheduled and coordinated through the Electronic Products Development Branch under Administrative and Customer Services Division.

For course specifics see Resources - "ICG Training Outline" on the ICG IntRAnet Page:
<http://cww.census.gov/tab/www/icg.htm>

5 Reference Materials and Related Standards

5.1 Reference Materials

Internet Dissemination, Coordination, & Management Page
<http://cww.census.gov/tab/www/icg.htm>

World Wide Web Consortium is the best place for information on HTML and Cascading Style Sheets (CSS).
<http://www.w3.org/>

World Wide Web Consortium's Validator Service verifies HTML code.
<http://validator.w3.org/>

Information on web-safe color palettes
<http://www.lynda.com/hex.html>

5.2 Related Standards

IT Standard 17.0.0: Design and Development of Accessible Software

IT Standard 18.0.0: Design and Development of Web-Based Applications

IT Standard 7.0.5: Personal Computer (PC) Hardware Configuration

IT Standard 4.0.1: Workstation Configuration

These standards are on the Census Bureau Bulletin Board under #IT Policies and Standards.

6 Conversion Requirements

A subteam (representing various divisions) from the Internet Coordinators Group developed these Internet document standards. The Internet Coordinators and Data Disseminators are currently following the requirements contained in IT Standard 15.0.0 dated June 1, 1998. Therefore, it should not be a problem to convert existing documents to meet the new standards contained in this version.

7 Compliance Methods and Timetable

7.1 Compliance Methods

The Internet Coordinators Group (ICG) is self-governing and expects compliance among its members. The ICG membership consists of either content developing divisions or, occasionally, content oversight areas. Both are administration groups and have one or more disseminators/members. One coordinator is assigned for each division/area with responsibilities of communications and oversight among 'local' membership. These coordinators **shall** participate in the ICG process.

Division chiefs **should** work with Internet Coordinators to develop Internet plans for their respective areas. Plans **shall** encompass the needs of the Bureau and take into consideration current communications and marketing efforts.

The Internet Management Team (IMT) is responsible for upholding compliance. Waivers to these standards **shall** be allowed if fully justified by processing requirements, security or cost constraints. All applications for waivers **shall** be made in writing, following the procedures established through the IT Standards and Uniform Products Program. In case of a standards dispute, the ICG subteam on standards would reconvene to recommend a resolution of the outstanding issue(s). Recommendations would be forwarded to both the division and the IMT for approval. The IMT would handle appeals from administration groups (divisions) before the release of a recommendation to the Standards Management Team who issues the final recommendation.

Administrative and Customer Services Division (ACSD) monitors and enforces standards compliance. In addition, ACSD retains the authority to request revocation of Internet write privileges of any Data Disseminator who refuses to comply with the standards contained in this document. The Systems Support Division retains the implementation of Internet write privileges.

7.2 Timetable

The Internet Document Standards are already in effect so changing to these new requirements should be an easy process.

Updates to these standards will occur at a one year interval. The Internet Coordinators Group and the Internet Management Team expect compliance on newly created pages within one month of each reissue date.

Under special circumstances, the Internet Management Team reserves the right to impose immediate changes to the standards as necessary.

8 Authority

8.1 Policy

Cited from Census Bureau Information Technology Policy Manual:

CHAPTER 4 - INFORMATION AND DATA MANAGEMENT

A. OVERVIEW

Modern IT continues to present the Bureau with opportunities to improve the management of its programs and the ability to serve the public more effectively.

4. Data Administration - Data definitions and formats, and their implementation, access, and documentation, are in a form that facilitates interchange with internal users and external customers. Data is managed so that it can be defined independently of the processes that create or use it, and so that it can be shared among many processes.
5. Data Documentation - Documentation for each data set or file is easily accessible and directly linked to the data it describes.

C. RESPONSIBILITIES

3. Data Management Compliance - Each Associate Director and Division and Office Chiefs ensures compliance with information and data management policies and administers data management within his or her directorate.
4. Evaluating New Technology - The Associate Director for Information Technology directs and coordinates the evaluation of new information management technologies to determine their potential applicability and effectiveness in Bureau information collection and dissemination activities.

6. Disseminating Information - The Associate Director for Information Technology establishes and maintains a technology infrastructure for effective and efficient dissemination of information.

8.2 Mandatory Standards

There are no mandatory standards.

Important Dates	
Originally Issued	June 1, 1998
Reissued	November 24, 1999
Next Review Scheduled	November 24, 2000
Compliance	see 7.2

Contact for Additional Information
<p>To find out:</p> <ul style="list-style-type: none">! where to send comments on specific standards! where to send suggestions to improve standards! what standards have been issued! general information about IT Standards and Uniform Products Program <p>Look at: IT Standards in Progress Bulletin Board IT Policies and Standards Bulletin Board</p> <p style="text-align: center;">or</p> <p>Contact: Software and Standards Management Branch in SSD 457-2142; Room 1385-3; 457-4606 (FAX) E-mail to <i>IT Standards Program</i> account</p>